PERFORMANCE AND CHARACTERIZATION OF U.S. NAVAL OBSERVATORY CLOCKS

Lee A. Breakiron and Demetrios N. Matsakis
Time Service Department
U.S. Naval Observatory
Washington, DC 20392, USA

Abstract

Six years of U.S. Naval Observatory clock data are analyzed to determine the optimal data filter length and the amplitude and frequency of occurrence of statistically significant changes in the frequencies and frequency drifts of cesium-beam frequency standards and hydrogen masers by means of relative and correlation-corrected N-cornered-hat analyses of the frequency stabilities of postprocessed mean timescales. The effects of temperature and humidity on cesium and maser frequencies and drifts are also investigated.

1 INTRODUCTION

The U.S. Naval Observatory (USNO) maintains an ensemble of about 50 HP5071 cesium-beam frequency standards and a dozen Datum-Sigma Tau hydrogen masers at its Washington, DC site, which are kept in environmentally controlled conditions. Since each of these clocks has been acquired, time (phase) differences between them and the Master Clock have been recorded once per hour, using a Data Acquisition System (DAS) of coaxial cables, multiplexed switches, time-interval counters, and computers.

One motivation for this work is simply to use our available data to determine the long-term, observed, statistical behavior of all the clocks in our ensemble. Although the lack of an absolute frequency reference makes it technically impossible to determine the absolute stability of any individual clock or ensemble, a considerable body of literature [e.g. 1-6] has provided some tools that allow mathematically precise "N-cornered-hat" estimates, subject to specific limitations. In this work we have chosen to use the tools of Torcaso et al. [3,4] because they were available for use in a form that allows simple batch processing of hundreds of clock combinations. The maser stabilities derived from this N-cornered hat analysis are shown to be a few tenths of a dB higher than stabilities derived from simpler analyses comparing individual clocks to USNO unsteered cesium and maser averages, EAL [7], and TT99 [8,9]. The reasons for this are discussed below.

A second motivation for this work has to do with the operational issues concerning the generation of USNO mean timescales. These timescales are actually integrated frequency scales, created by averaging clock frequencies that have been detrended by removing clock-specific rates (frequencies) and drifts (linear change of rates with time). The timescales differ in how they weight clocks, but all clocks are currently characterized through comparison with an unsteered cesium mean timescale [10,11]. Since all frequency standards exhibit nonwhite

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **NOV 2000**	2. REPORT TYPE	3. DATES COVERED **00-00-2000 to 00-00-2000**

4. TITLE AND SUBTITLE **Performance and Characterization of U.S. Naval Observatory Clocks**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Naval Observatory,Time Service Department,Washington,DC,20392**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM001509. 32nd Annual Precise Time and Time Interval (PTTI) Meeting, 28-30 Nov 2000, Reston, VA

14. ABSTRACT
Six years of U.S. Naval Observatory clock data are analyzed to determine the optimal data filter length and the amplitude and frequency of occurrence of statistically significant changes in the frequencies and frequency drifts of cesium-beam frequency standards and hydrogen masers by means of relative and correlation-corrected N-cornered-hat analyses of the frequency stabilities of postprocessed mean timescales. The effects of temperature and humidity on cesium and maser frequencies and drifls are also investigated.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**19**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

noise over long enough periods, a problem develops because there is no optimal manner to average nonwhitened noise. In practice, USNO has taken advantage of the large number of clocks in its ensemble by using the timescale itself as a reference with which to determine epochs of significant changes in clock rates and drifts. The epochs of these "clock breaks" are chosen by inspection of the whiteness of the frequency data and linearity of the phase data.

These breaks may occur at any time due to changes in environmental conditions or spontaneous changes within the clock. Such breaks are checked for during each hourly measurement and computation of the real-time mean timescale, and the clocks involved are deweighted. The clocks are later reweighted after sufficient data have accumulated to accurately model the clocks' rates and (if any) drifts. Clocks are also checked for, and deweighted in the event of, general degradations in stability.

Cesiums are weighted equally in the timescale computation due to lack of evidence that any other procedure is optimal, at least using weights based on Allan variances [11,12] or rate solution variances. This is probably due to the deweighting of any clocks showing changes in performance or character. In the combined cesium-maser timescale toward which the Master Clock is steered, masers are deweighted relative to cesiums in a manner varying with time since the latest measurement [10]. Clock parameter recharacterization is done prior to weighting and is relative to the mean of the other, weighted clocks. The number of clocks recharacterized at any one time is limited to a small fraction of the weighted ensemble, lest the stability of the whole timescale be afflicted by errors associated with indeterminacy because of the lack of an absolute reference.

How well the rates and drifts of the individual clocks are modelled and how quickly any changes therein are recognized and allowed for have definite effects on the stability of the mean timescales, both those that are used diagnostically in postprocessing and those that are used in real time as a target toward which to steer the USNO Master Clock. The stability of the Master Clock is, however, not strongly affected by the rate of clock recharacterization, since even the most aggressive recharacterization studied herein does not result in any significant destabilization over the monthly intervals between steers of the mean timescale toward TAI. This study is the first of several different empirical and theoretical attempts at USNO to quantify the costs and benefits of different levels of precision in clock characterization. Another objective is the determination of the average size of and average interval between clock breaks, i.e. the temporal stability of these frequency standards.

CLOCK DATA AND ANALYSIS

Data reported here were taken from MJD 49532-51857 (29 Jun 94-9 Nov 00), using a DAS which each hour switches each clock's 5 MHz signal to a time-interval counter for a phase measurement against the Master Clock. These measurements have an hourly precision of about 50 ps rms, but the very long-term accuracy is 1 ns peak to peak. Although recently experimental, lower-noise measurement systems are being used in parallel [13], DAS switch data are used throughout this analysis.

The physical environment of the USNO clocks are usually maintained to an rms of 0.1°C and 1% relative humidity through the use of temperature-stabilized chambers inside of temperature-controlled rooms. However, there are short-term fluctuations in any given chamber of a degree or more, due to equipment failure about once a year. Also, each clock can experience long-term temperature changes when nearby equipment is relocated or when components of the environmental control system are replaced through normal maintenance. Humidity variations

are usually associated with temperature variations, but no significant clock variations have been ascribed purely to humidity variations. Temperature variations, however, can noticeably affect maser frequencies, whose temperature dependence has been reported to be up to $9 \cdot 10^{-15}/°C$ [14] in absolute value. As part of this work, we have also examined available "health and status" information from the RS-232 ports of the masers and cesiums, which are used here as diagnostic tools [15,16]. We have found, as did Chadsey [17] for cesiums, that neither maser nor cesium frequency variations correlate well with the health and status information, except for hardware (e.g. power supply) failures and maser temperature variations.

Rates are determined by averaging the first differences of the hourly phase measurements, unless there appears to be significant drift, in which case rate and drift are derived from a linear least-squares fit to the first differences. This procedure has been found to be optimal for our data [18], because a sampling time of 1 hour is in the white FM noise region of both the cesiums and our DAS measurement system (hence, the masers as well). While rate can be inferred from initial and final phase measurements, the above procedure permits data editing and confidence level determination.

The least precise aspect of USNO clock characterization is the determination of the clock breaks. In practice, the methods have varied considerably over the years. In an on-line situation, it has not been unusual for the data analysts to remove a clock from the averaging simply because a few hours' data show a trend that could, over a weekend, affect the mean timescale significantly. This is often a prudent decision, because the loss of just one good clock affects stability by the square root of N, while a mischaracterized clock affects the average by 1/N, and because in general robustness is more critical than stability. Another reason to err on the side of caution is because USNO Master Clock steering strategies are designed to protect against short-term mean instabilities, but long-term trends are removed only by the steers to UTC (BIPM) [19,20].

Even in postprocessed work such as this, the difference between a time series characterized by clock breaks and one characterized by flicker noise can be as much philosophy as substance [21]. For this work, clock breaks were determined by four different sets of criteria, characterized by different levels of increasingly strict modelling. All methods discarded obviously bad data, and inserted clock breaks at times of disturbances and repairs, such as clock moves and beam tube replacements. To these, all methods added additional clock breaks of a number proportional to the level of significance adopted. These criteria are summarized in Table 1. The strictest criteria (Set #4) are those currently used in day-to-day timescale operations, though these are approximate; the actual values varied with the empirical philosophy of the data analyst at the

Table 1. Minimum Clock Break Levels (Absolute Values)

	Cesiums		Masers	
Set #	Rate Change	Drift Change	Rate Change	Drift Change
1	$5.0 \cdot 10^{-14}$	$5.0 \cdot 10^{-16}/day$	$3.0 \cdot 10^{-14}$	$1.0 \cdot 10^{-16}/day$
2	$1.5 \cdot 10^{-14}$	$2.5 \cdot 10^{-16}/day$	$4.0 \cdot 10^{-15}$	$1.5 \cdot 10^{-17}/day$
3	$4.6 \cdot 10^{-15}$	$1.0 \cdot 10^{-16}/day$	$9.3 \cdot 10^{-16}$	$4.6 \cdot 10^{-18}/day$
4	$3.5 \cdot 10^{-15}$	$5.8 \cdot 10^{-17}/day$	$4.6 \cdot 10^{-16}$	$3.5 \cdot 10^{-18}/day$

time the data were current. Moreover, the tabulated values do not reflect the entire story, since there is some variation with effective filter size. For criteria Sets #1-3, clock breaks would be assumed if changes smaller than the tabulated values occur over long periods of time.

By way of illustration, Figure 1 shows how two different cesium clocks (serial numbers 114 and 1097) were broken into clock breaks by the four different sets of criteria. Also, Table 2 summarizes the total number of clock breaks detected.

Table 2. Number of Clock Breaks Detected

Set #	Cesiums	Masers	Totals
1	150	39	189
2	260	90	350
3	312	92	404
4	577	155	732

POSTPROCESSED TIME SCALES

In order to compare the four different clock-characterization methods, postprocessed timescales were generated using a modified form of Percival's [22] algorithm through the in-house least-squares program SuperP [12]. Figure 2 compares cesium-only mean timescales, generated by SuperP with TT99 as the reference. Similar results are found with maser-only mean timescales or using the EAL as a reference. It appears that the USNO data have in the past been over-aggressively characterized, although the least aggressive set of criteria was also suboptimal. This conclusion is supported by analysis of timescales generated using SuperP on different halves and thirds of the USNO ensemble by performing N-cornered-hat frequency stability analyses upon the independent thirds and computing the Hadamard deviations [6].

For the rest of this work, clock characterizations will be determined using Set #3. It should be noted that Set #2 may be equally valid, and would result in slightly higher stability measures, because it has clock breaks at fewer positions where the data begin to look nonwhite. No matter which criteria set was used, the clocks on the whole had indistinguishable stability measures for sampling times τ of < 10 days, because the clock weighting and clock-break determinations were made on a longer scale than this. For $\tau > 10$ days, the computed cesium and maser stabilities tended to be about 0-1 and 2 dB higher, respectively, for Set #1 compared with the other sets.

STATISTICAL MEASURES AND REFERENCE TIMESCALES

The stability characteristics of individual clocks were compared to the following four available references: USNO's unsteered maser mean timescale (MM), USNO's unsteered cesium mean timescale (CM), EAL, and TT99. Figure 3 compares Allan deviations [1] of the MM with each of the other three references using data corrected for drift. Note that the MM is considerably more stable than the CM in the short term, and its stability approaches that of the CM in the long term because both cesiums and masers are detrended against the same standard.

Figure 4 shows the Allan and Hadamard deviations for the median of the masers and cesiums relative to the MM, and for the mean of the best third of the masers and cesiums relative to the maser mean. The stabilities are consistent with similar analyses wherein the reference was the cesium mean, EAL, or TT99. Figure 4 implies that at least 30 days of data are required to determine a clock's frequency to within 4 parts in 10^{15}, and that at least 60 days of data are required to determine a clock's drift to within a part in 10^{16}/day. This is consistent with established USNO practice to not weight a clock until it has at least 30 days of data that are stable in rate and display no significant drift, or 60 days of data that are stable in drift and, aside from that, rate.

Figure 5 plots the N-cornered-hat stabilities for all the aforementioned combinations, as well as drift removal, derived using the methods of Torcaso et al. [3,4]. The derived stabilities agree for the Hadamard deviations of the masers and for both Allan and Hadamard deviations of the cesiums. The high values for the Allan deviations of our masers are due to the sensitivity of the N-cornered-hat method of Torcaso et al. to noisy clocks, such as masers with large drifts. Such high values do not appear in our plots of Allan deviation for drift-corrected data. We note also that the technique failed to give reasonable values for masers in hybrid (maser and cesium) ensembles.

Note that the manufacturer's rating for a typical USNO maser's time deviation (TDEV) is 100 ps over a day. Since this is comparable to the noise in the measurement system that produced the data reported here, we do not report maser Allan deviations for $\tau < 2$ days or Hadamard variances for $\tau < 4$ days.

CESIUM-BEAM CLOCK PERFORMANCE

High-performance HP5071A cesium-beam frequency standards have been in general use at USNO since 1993 and were the basis of a stability analysis in 1994 [23]. The lifetime of a typical high-performance tube, warrantied for 3 years, has been found to range from 5 to 7 years [17,24]. An initial "burn-in" phase of instability ranges from 0 to about 90 days, an example of which is shown in Figure 6a (relative to the MM). Figure 7 plots the length of each burn-in period, in days, vs. the MJD the tube was received at USNO.

An end-of-life period of high instability generally ranges from about 1 day up to about a year during which the instability is significantly higher or more nonwhite than normal. Figure 6b shows a typical tube demise, which is first signaled by a sharp rise in the electron multiplier voltage to its maximum voltage (2553 V), followed by a rise of the signal gain to its maximum (100%) [15]. A very slight increase in noise is observed up to 300 days before tube failures in about a quarter of the cases (e.g. Figure 6c). This is roughly consistent with Allan and Hadamard deviation statistics computed over each 160-day interval, and which show a very small degradation in the average stability as a function of the age of tube (perhaps 1 dB over the tube life).

Examining Set #3 of times of rate and drift changes, it was noted that significant changes in rate (more than about 3 parts in 10^{15} occur on the average of once every 312±15 days, the absolute value of the average rate change being 1.7±0.17 parts in 10^{14}. Despite the occurrence of very transient temperature excursions of about 10°C and changes of a few degrees over several hours in the environmental chambers, not a single significant frequent step could be unequivocally attributed to temperature or humidity change. This is consistent with the specifications and test results for these clocks [25].

Figure 8 plots the observed cesium drifts as a function of MJD, for Set #2 data. The increase in

drift scatter at the plot's extremities may be due, in whole or part, to insufficient measurement time.

As noted in the previous section, 60 days appears to be be the minimum length of an optimal filter for the determination of cesium rates. But 60 days also approximates the maximum length, as is evident in Figures 2-5 and 9. The improvement in rate accuracy slows down after 2 months because of nonwhite noise. Thus, increasing the filter length much beyond 2 months would subject the rate determinations to random walk FM.

HYDROGEN MASER PERFORMANCE

The frequency performance of six of our earlier Datum-Sigma Tau auto-tuned masers has been published in [23]. The relative sizes of our masers' drifts, as well as their associated errors, are plotted in Figure 10. As it is for the rates of cesiums, 60 days also appears to be the approximate upper limit on the optimal filter length for the drifts of masers (see Figures 3-5 and 11).

Examining Set #3 of times of rate and drift changes, it was noted that significant changes in drift ($> 5 \cdot 10^{18}$/day) occur on the average of once every 209±16 days, the absolute value of the average drift change being $7.3 \pm 1.7 \cdot 10^{17}$/day. About a fifth of the rate changes could be attributed to temperature excursions in the environmental chambers. Regarding drift, the three correlations found all concerned the same maser (NAV8). Spikes greater than 3°C or variations of 0.°5C prolonged over several hours could cause rate changes, though excursions as large as 8°C could have no effect, and only about fifth of temperature changes > 3°C caused rate changes.

Determination of reliable temperature coefficients from these data is problematic because it is based on failures or adjustments of the environmental chambers, rather than on controlled experiments. Separate determination of the relative or absolute humidity coefficients is further complicated by the high correlation between relative humidity and temperature, and the fact that both tend to vary when an environmental chamber fails.

Temperature-induced frequency variations were not always consistent for the same maser. Short-term correlations were observed ranging from approximately -1.8 to +1.5 parts in 10^{14}/°C. Figure 12 shows long-term variations for maser NAV3 indicative of a temperature coefficient on the order of of $-1 \cdot 10^{-14}$/°C. Parker [14] found long-term coefficients ranging from -9 to +1.3 parts in 10^{15}/°C.

As stated, we also found probable effects on drift. Figure 13 displays a long-term variation for maser NAV8 consistent with a temperature coefficient of $+0.81 \cdot 10^{-17}$/day/°C. Short-term correlations were found ranging from -0.7 to +5.8 parts in 10^{17}/day/°C.

Frequency variations in the masers often correspond to fluctuations in the maser health and status information, particularly the "top plate heater" voltage, the changes in which are related both to chamber temperature and temperature gradients within the maser. Moreover, the measured temperatures are dependent on the placement of the sensors within the chamber. Hence, the aforementioned temperature coefficients are not necessarily reflective of the temperature experienced by the maser.

CONCLUSIONS

Allan and Hadamard deviations of the USNO ensemble of HP5071 cesium standards and hydrogen masers were computed several different ways. Sixty days appears to be the optimal length for a filter on hourly first differences in the determination of clock rates and drifts. Thus, less aggressive recharacterization than is presently used operationally will improve USNO timescale stability, although it will not necessarily affect the short-term stability of the Master Clock over periods of less than 60 days.

Comparison of postprocessed timescales indicate that cesium rates and drifts should be recharacterized after changes of about 5 parts in 10^{15} and 1 part in 10^{16}/day respectively. Such changes occur about every 312 days, averaging 1.7 parts in 10^{14}. There appears to be an increase in the drift of new cesium tubes, but this may be due to insufficient averaging time.

Similar comparisons indicate that maser rates and drifts should be recharacterized after changes of about 1 part in 10^{15} and 5 parts in 10^{18}/day respectively. Such changes occur about every 209 days, averaging 7.3 parts in 10^{17}/day.

While no frequency dependence on temperature or humidity could be ascertained for the cesium standards, the masers evinced frequency dependences of 1 or 2 parts in 10^{14}/°C in absolute value and drift dependences of about 1 to a few parts in 10^{17}/day/°C in absolute value.

DISCLAIMER

References to specific commercial products do not imply an endorsement by the U.S. Naval Observatory. Although the analysis of clock stabilities is believed to be accurate with regard to the actual experience of the USNO, it should not be construed that this would be characteristic of clocks maintained at other laboratories at the same period, nor of clocks currently marketed by any manufacturer.

ACKNOWLEDGMENTS

We would like to thank the past and current staff of the USNO Time Service Department for their dedicated efforts on behalf of the Master Clock, and in particular Harold Chadsey, Randy Clarke, and Tony Kubik. We also thank Fred Torcaso for advice on N-cornered-hat solutions, and Bill Riley (Datum) for his computer program Stable used to check in-house statistical software, and Ken Senior for assistance with the figures.

REFERENCES

[1] D. Allan 1987, *"Time and frequency (time-domain) characterization, estimation, and prediction of precision clocks and oscillators,"* **IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, UFFC-34,** 647-654.

[2] P. Tavella, and A. Premoli 1993/1994, *"Estimating the instabilities of N clocks by measuring differences of their readings,"* Metrologia, 479-486.

[3] F. Torcaso, C. R. Ekstrom, E. A. Burt, and D. N. Matsakis 1999, *"Estimating frequency stability and cross-correlations,"* Proceedings of the 30th Annual Precise Time and

Time Interval (PTTI) Systems and Applications Meeting, 1-3 December 1998, Reston, Virginia, USA, pp. 69-81.

[4] C. R. Ekstrom, F. Torcaso, E. A. Burt, and D. N. Matsakis 1999, *"Estimating the stability of N clocks with correlations,"* Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum (EFTF) and the IEEE International Frequency Control Symposium, IEEE Publication 99CH36313, 13-16 April 1999, Besançon, France, pp. 168-172.

[5] F. J. Galindo, and J. Palacio 2000, *"Estimating the instabilities of N correlated clocks,"* Proceedings of the 31st Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 7-9 December 1999, Dana Point, California, USA, pp. 285-296.

[6] S. T. Hutsell 1996, *"Relating the Hadamard variance to MCS Kalman filter clock estimation,"* Proceedings of the 27th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 29 November-1 December 1995, San Diego, California, USA, pp. 291-301.

[7] C. Thomas, P. Wolf, and P. Tavella 1994, *"Time scales,"* BIPM Monograph 94/1.

[8] B. Guinot 1995, *Scales of time*, **Metrologia, 31,** 431-440.

[9] TT99 from BIPM by anonymous ftp at 62.161.69.5/pub/tai/scale/ttbipm.99

[10] L. A. Breakiron 1992, *"Timescale algorithms combining cesium clocks and hydrogen masers,"* Proceedings of the 23rd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 3-5 December 1991, Pasadena, California, USA, pp. 297-305.

[11] L. A. Breakiron 1989, *"The effects of data processing and environmental conditions on the accuracy of the USNO timescale,"* Proceedings of the 20th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 29 November-1 December 1988, Tysons Corner/Vienna, Virginia, USA, pp. 221-235 (see errata in [10], p. 302).

[12] D. Matsakis, and L. A. Breakiron 1999, *"Postprocessed timescales at the U. S. Naval Observatory,"* Proceedings of the 30th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 1-3 December 1998, Reston, Virginia, USA, pp. 19-33.

[13] D. N. Matsakis 2000, *"Recent and pending improvements at the U.S. Naval Observatory,"* Proceedings of the 31st Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 7-9 December 1999, Dana Point, California, USA, pp. 257-265.

[14] Parker, T. E. 1998, *"Influence of environmental factors on hydrogen maser frequency stability,"* Proceedings of the 12th European Frequency and Time Forum (EFTF), 10-12 March 1998, Warsaw, Poland, pp. 126-131.

[15] H. Chadsey, and A. Kubik 1998, *"Maintenance of HP 5071A primary frequency standards at USNO"* Proceedings of the 29th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 2-4 December 1997, Long Beach, California, USA, pp. 49-59.

[16] A. Kubik 2000, private communication

[17] H. Chadsey 2000, *"An automated alarm program for HP5071A frequency standards,"* Proceedings of the 31st Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 7-9 December 1999, Dana Point, California, USA, pp. 649-655.

[18] L. A. Breakiron 1994, *"A comparative study of clock rate and drift estimation,"* Proceedings of the 25th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 29 November-2 December 1993, Marina del Rey, California, USA, pp. 401-411.

[19] L. A. Breakiron 1996, *"Frequency steering of hydrogen masers,"* Proceedings of the 1996 IEEE International Frequency Control Symposium, IEEE Publication 96CH35935, 5-7 June 1996, Honolulu, Hawaii, USA, pp. 1113-1122.

[20] P. A. Koppang, and D. N. Matsakis 2000, *"New steering strategies for the USNO Master Clock,"* Proceedings of the 31st Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 7-9 December 1999, Dana Point, California, USA, pp. 277-284.

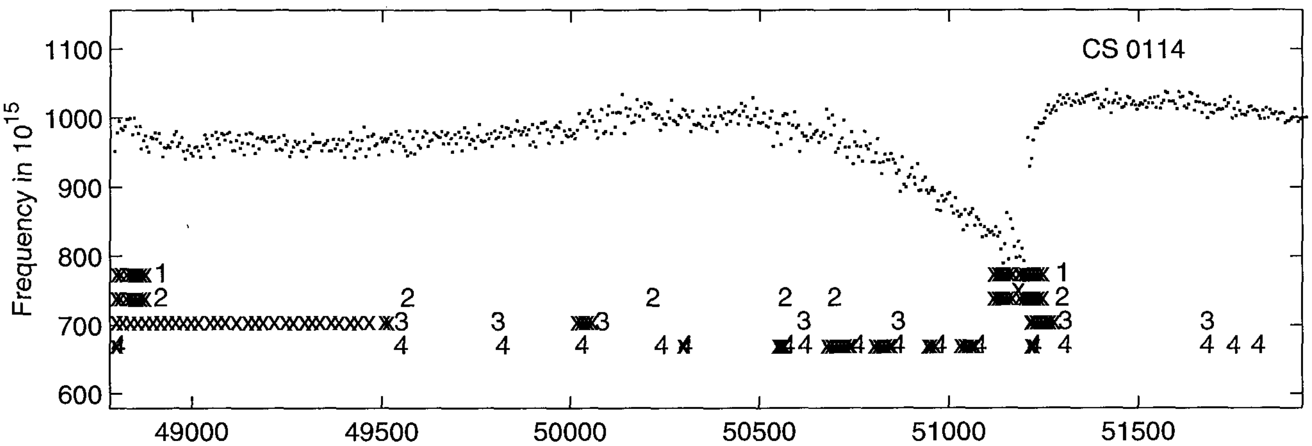
[21] G. J. Dick, and C. A. Greenhall, *"L.O. limited frequency stability for passive atomic frequency standards using "square wave" frequency modulation,"* Proceedings of the 1998 IEEE International Frequency Control Symposium, IEEE Publication 98CH36165, 27-29 May 1998, Pasadena, California, USA, pp. 99-103.

[22] D. B. Percival 1978, *"The U.S. Naval Observatory clock time scales,"* **IEEE Transactions on Intrumentation and Measurement, IM-27**, 376-385.

[23] L. A. Breakiron 1995, *"The effect of clock errors on timescale stability,"* Proceedings of the 26th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 6-8 December 1994, Reston, Virginia, USA, pp. 369-380 (see Errata vol. #2 for figures).

[24] J. A. Kusters, L. S. Cutler, and E. D. Powers 1999, *"Long-term experience with cesium beam frequency standards,"*, IEEE Publication 99CH36313, Proceedings of the 1999 Joint Meeting of the European Time and Frequency Forum (EFTF) and the IEEE International Frequency Control Symposium, 13-16 April 1999, Besançon, France, pp. 159-163.

[25] J. A. Kusters and J. L. Johnson 1992, *"A new cesium beam frequency standard performance data,"* IEEE Publication 92CH3083-3, Proceedings of the 1992 IEEE Frequency Control Symposium, 27-29 May 1992, Hershey, Pennsylvania, USA, pp. 143-156.



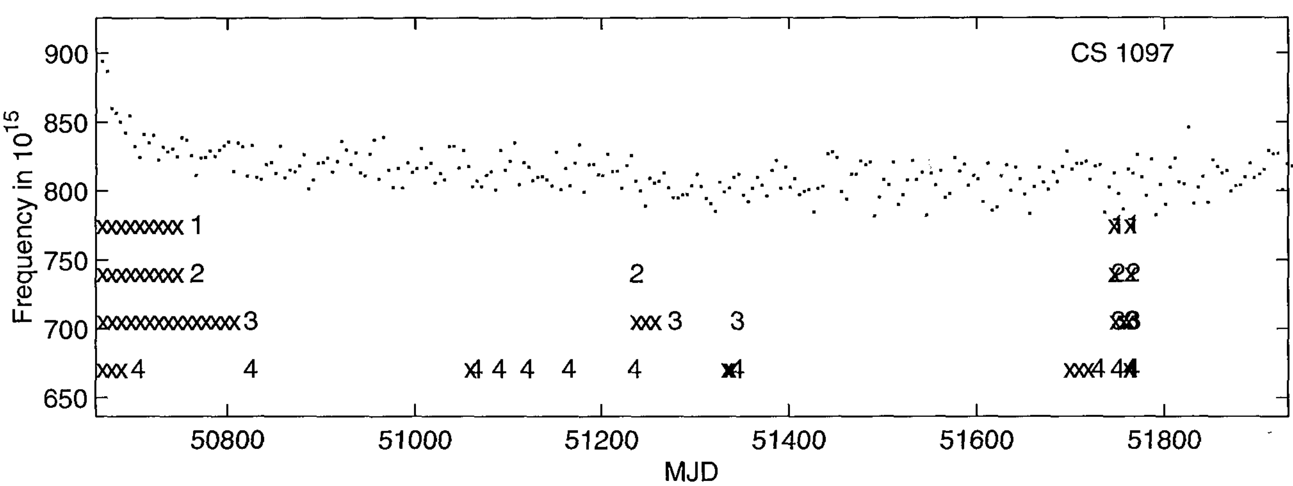


Figure 1. The frequencies of two cesium standards referenced to the cesium mean timescale, with the times of clock breaks determined by the four different sets of criteria indicated by "1," "2," "3," or "4." "x"'s in the row dedicated to a particular criteria set indicates that data were deweighted from that point on until the next clock break.

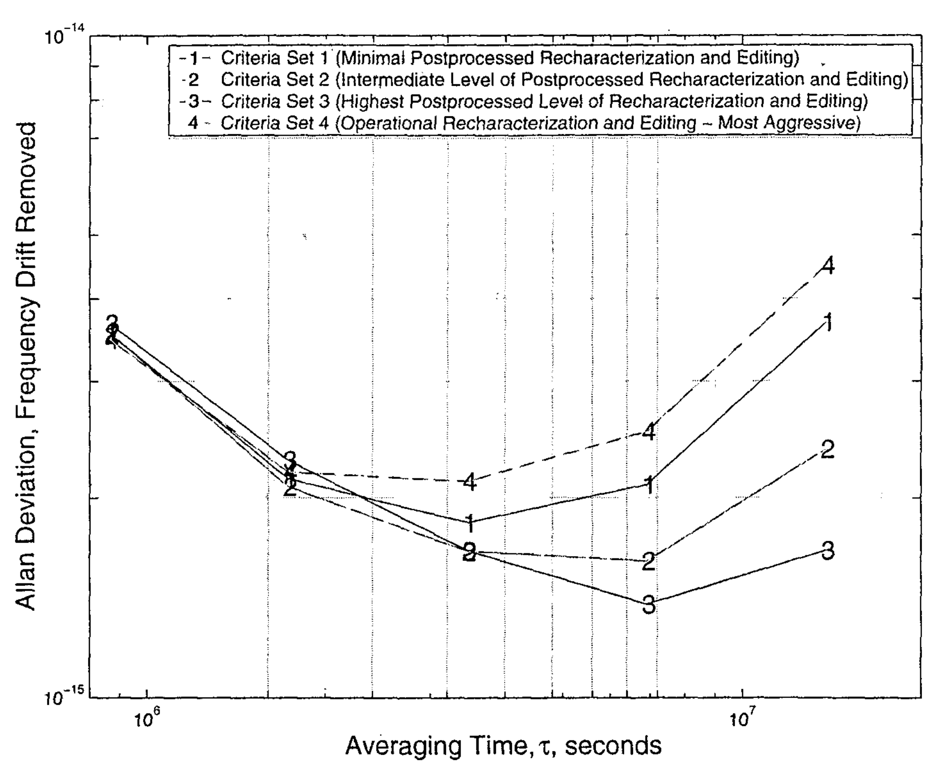


Figure 2. USNO unsteered cesium mean timescale frequency stabilities referenced to TT99.

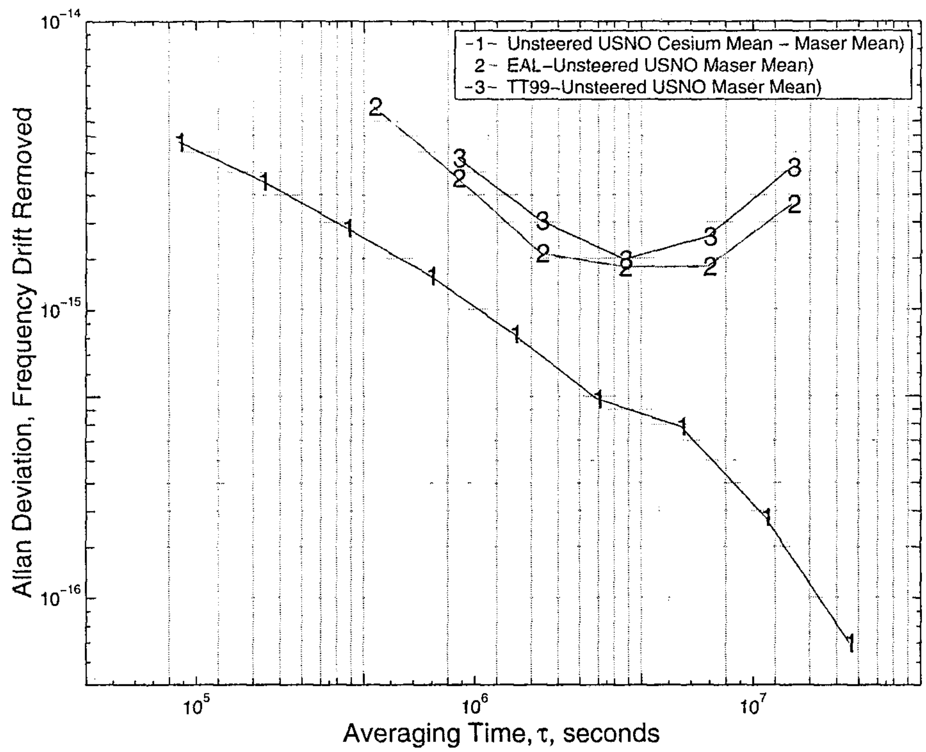


Figure 3. Timescale frequency stabilities referenced to the unsteered USNO maser mean timescale.

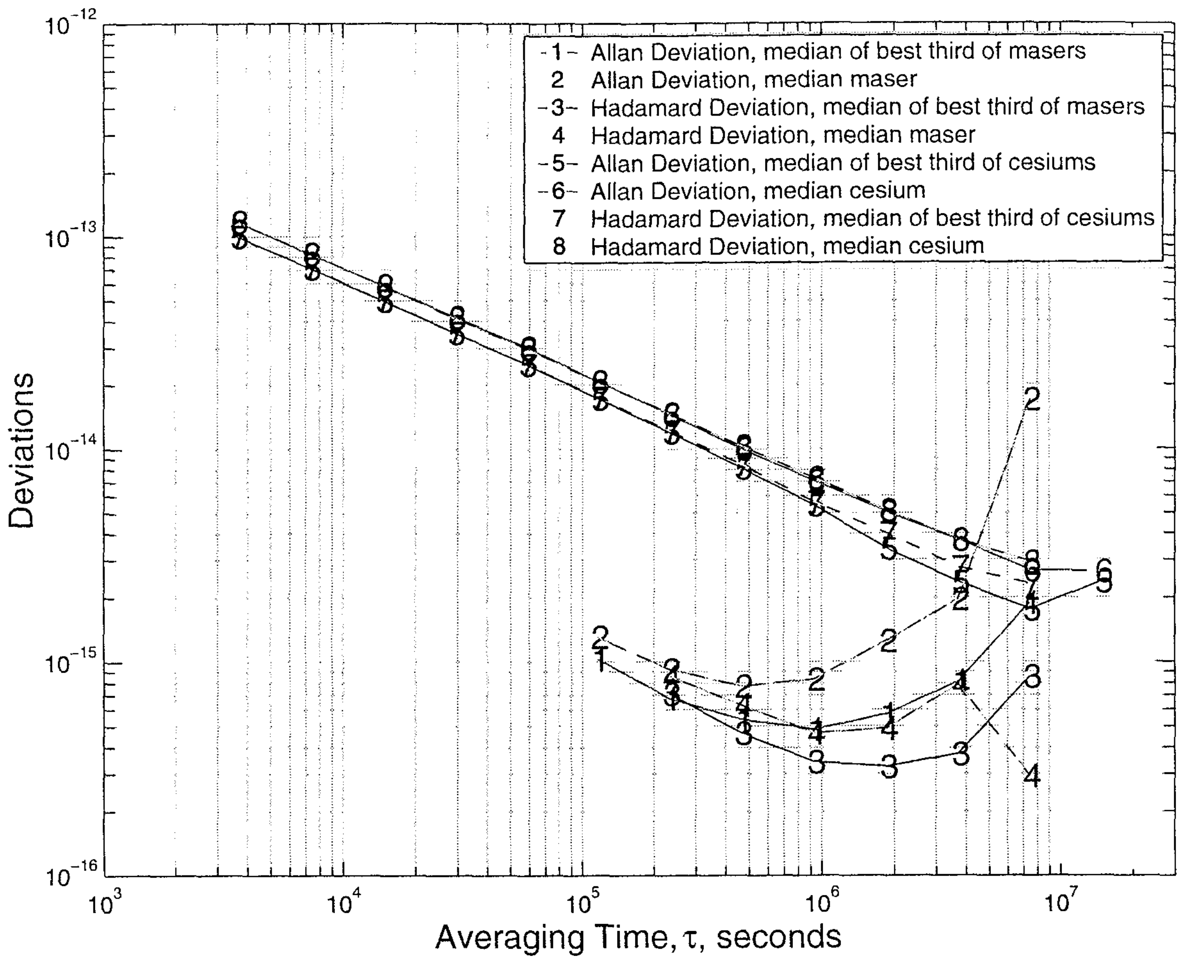


Figure 4. Postprocessed timescale frequency stabilities referenced to the unsteered USNO maser mean timescale using editing criteria Set #3.

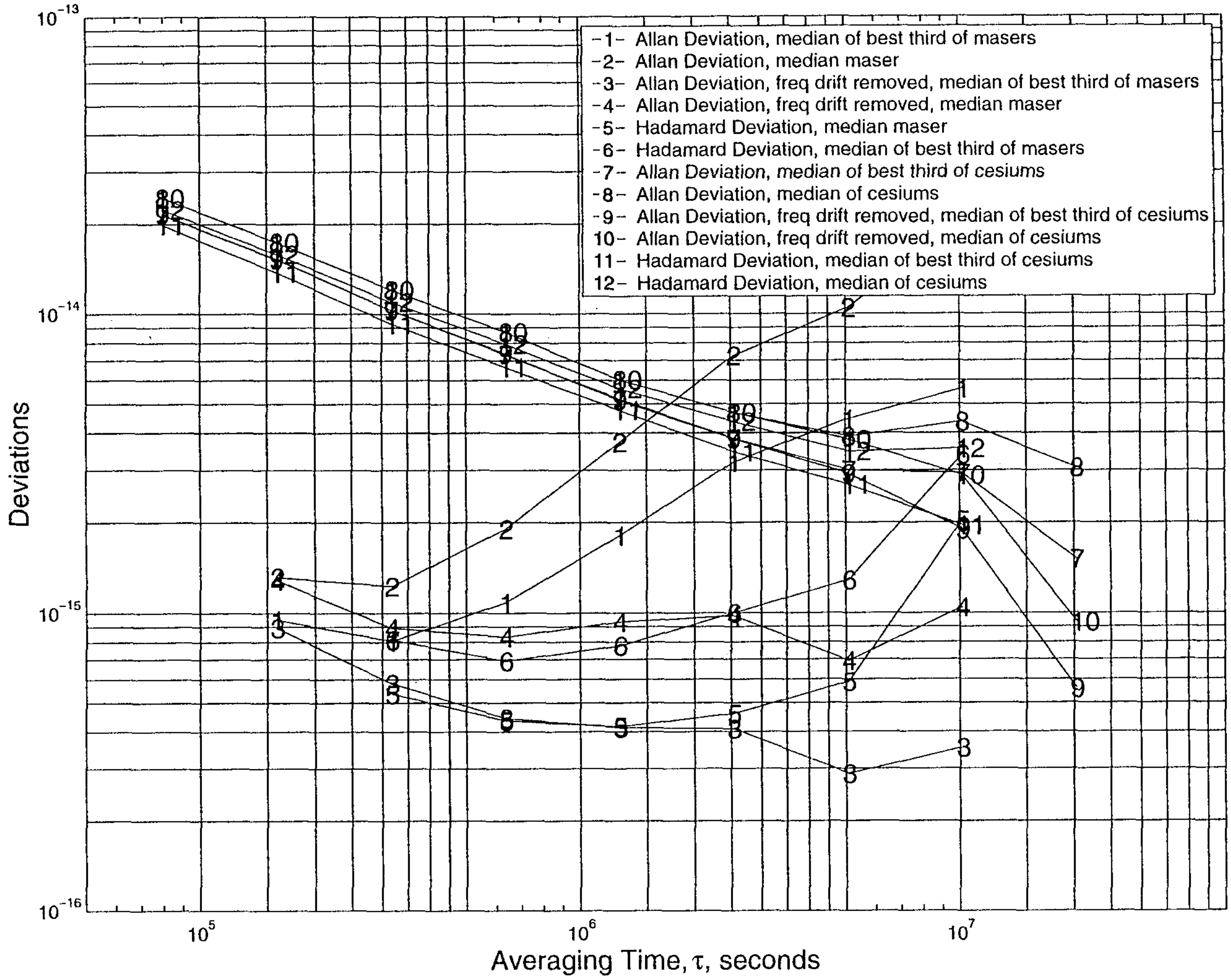


Figure 5. Postprocessed timescale frequency stabilities from N-cornered-hat analyses using editing criteria Set #3. The high values for curves "1" and "2" are due to the analyses systematically overstating the uncertainties of the lower-sigma clocks.

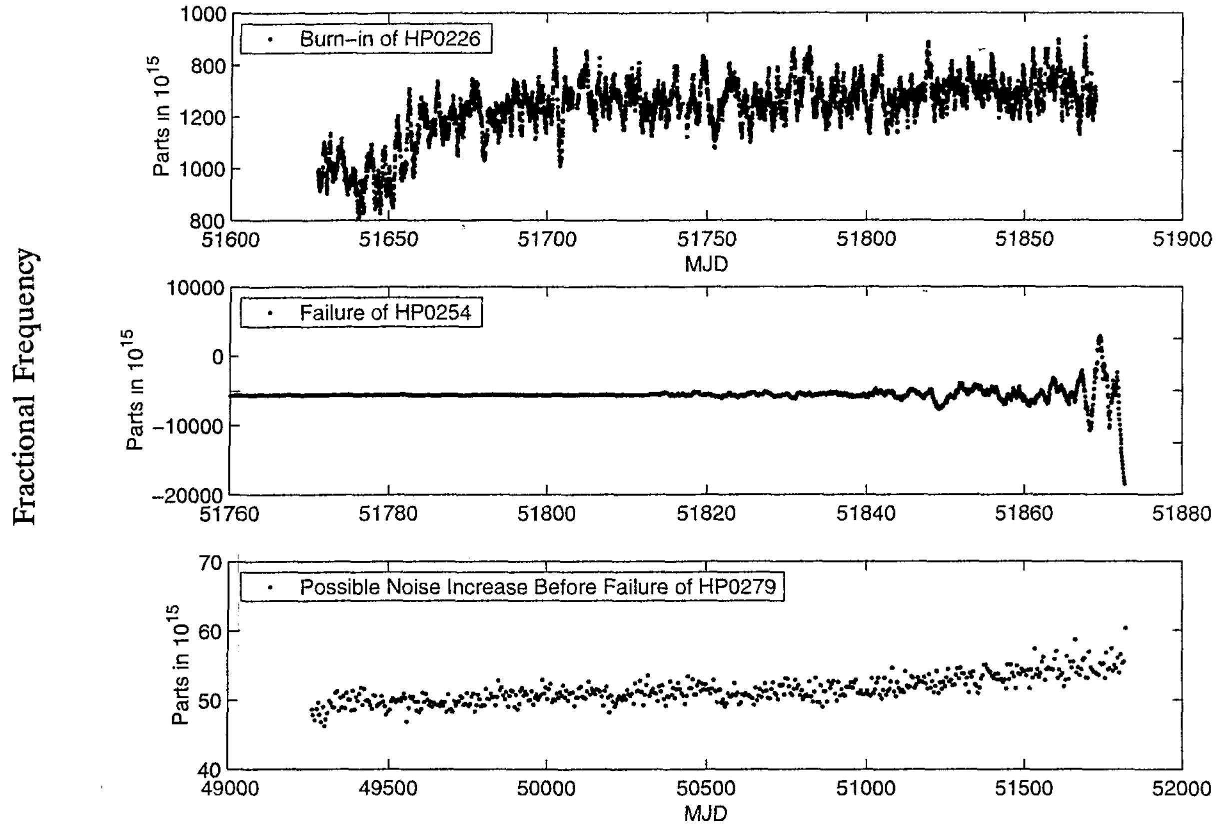


Figure 6. Samples of cesium beam tube burn-in and failure modes.

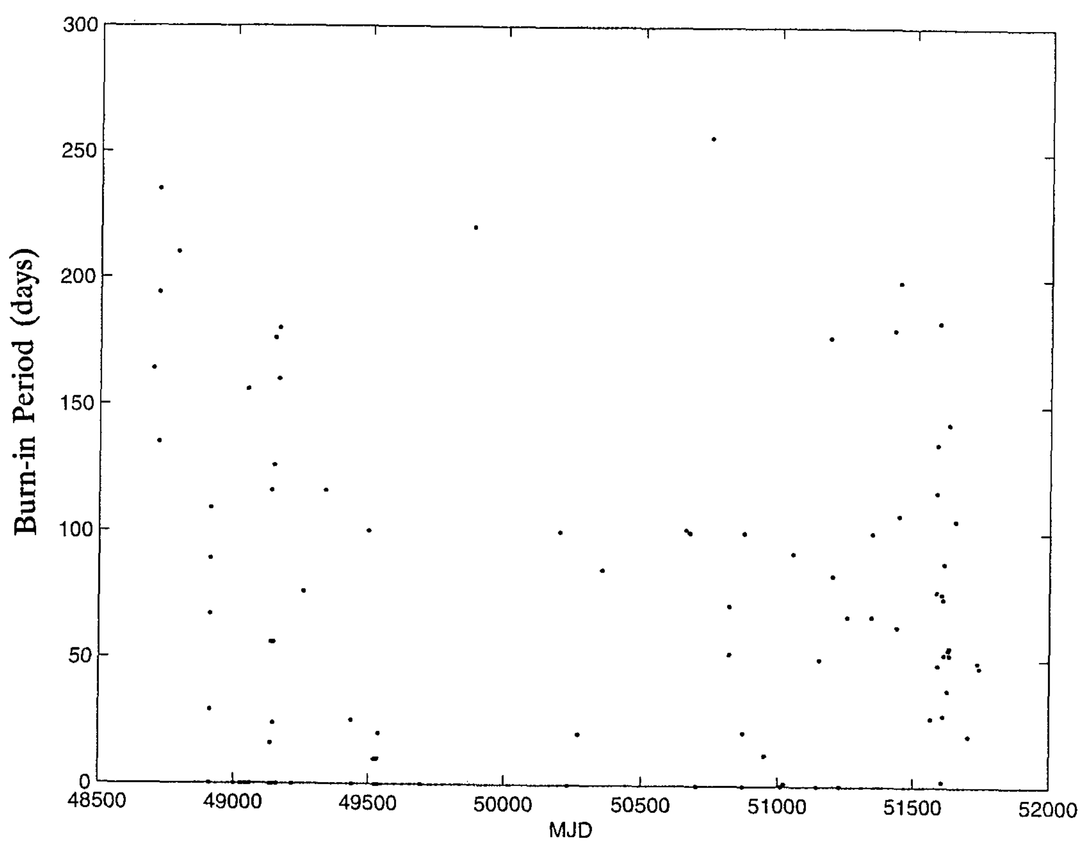


Figure 7. Length of cesium tube burn-in as a function of MJD.

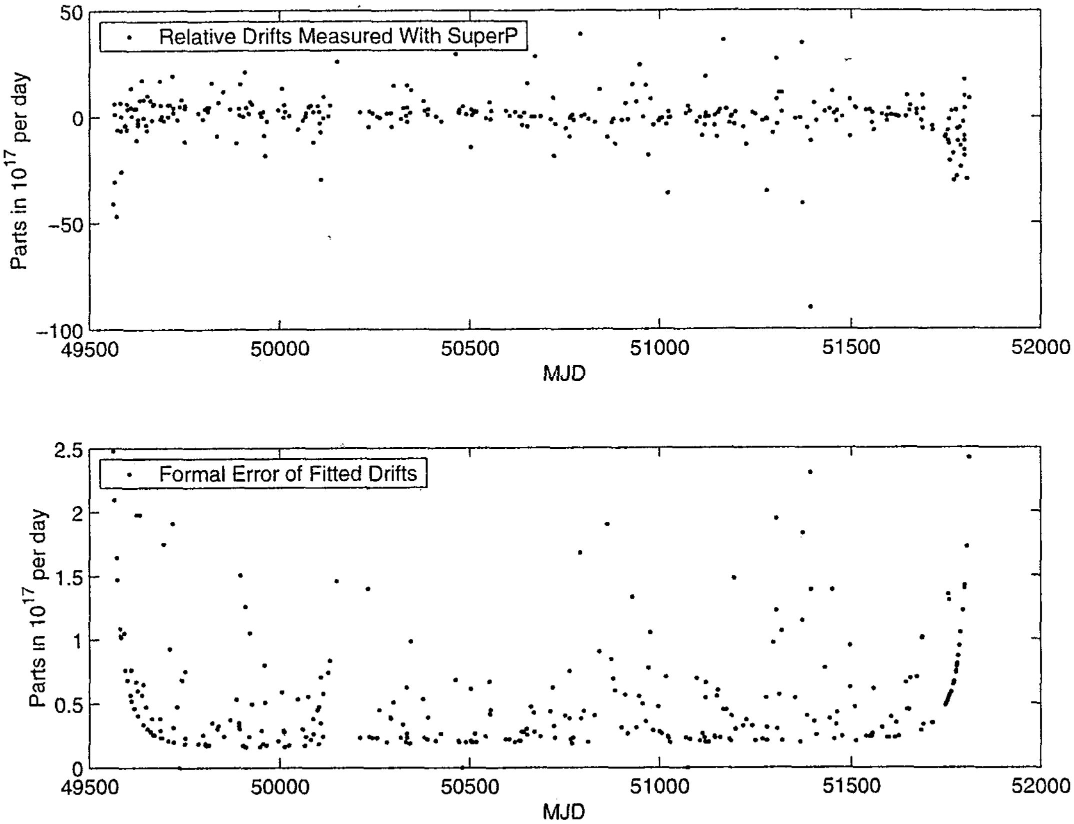


Figure 8. Fitted cesium clock drifts and associated standard errors as a function of characterization date.

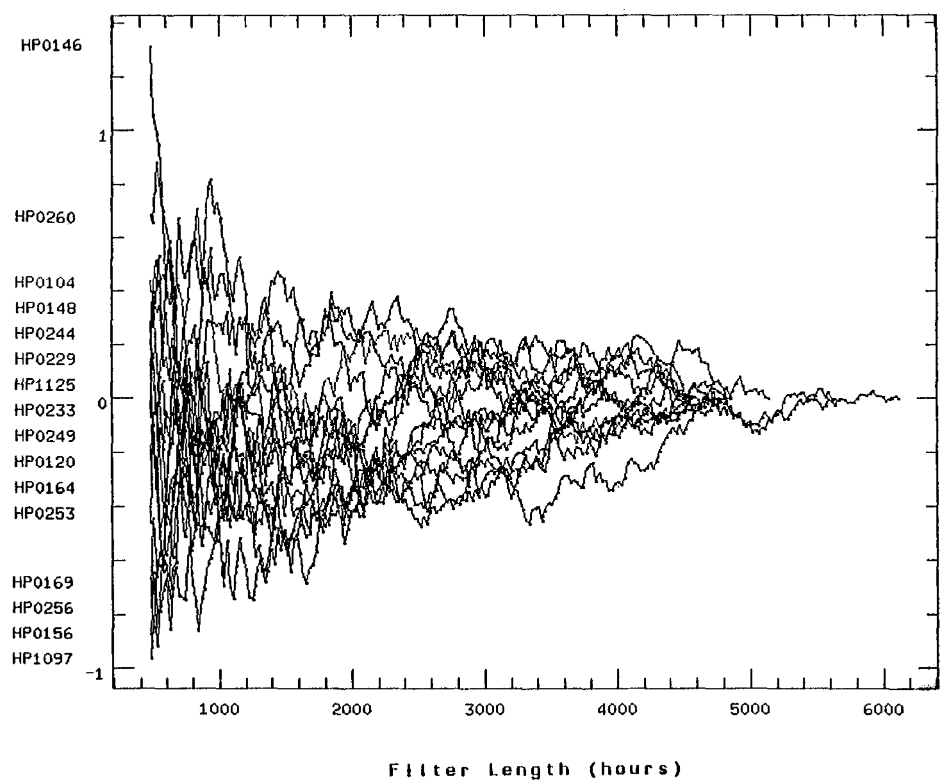


Figure 9. The decrease in the change of mean rate as data accumulate with time (converging toward zero at the final adopted rate) for a selection of 16 cesiums.

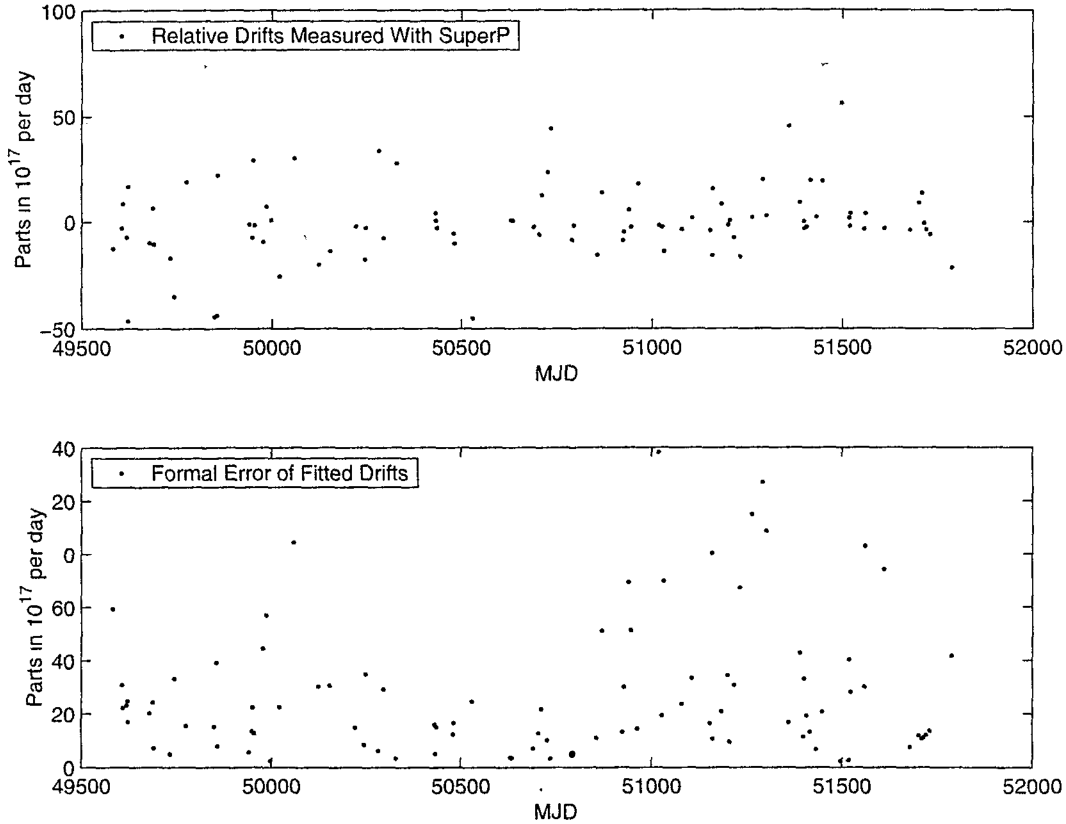


Figure 10. Fitted maser drifts and associated standard errors as a function of characterization date.

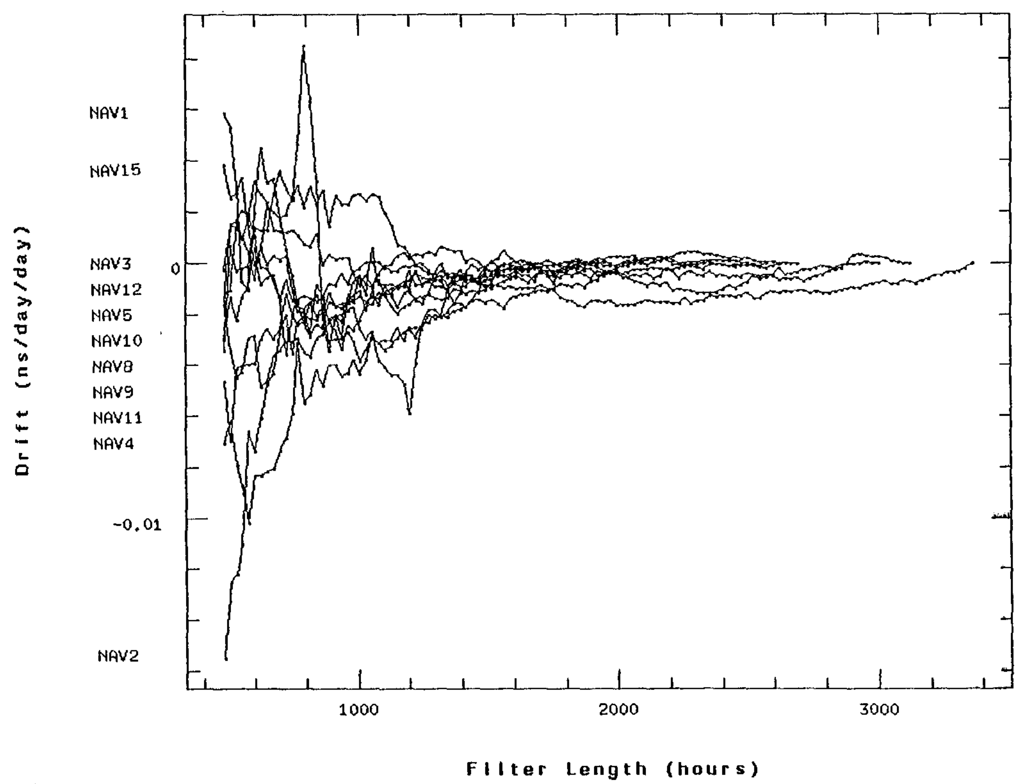


Figure 11. The decrease in the change of mean drift as data accumulate with time (converging toward zero at the final adopted drift) for 11 masers.

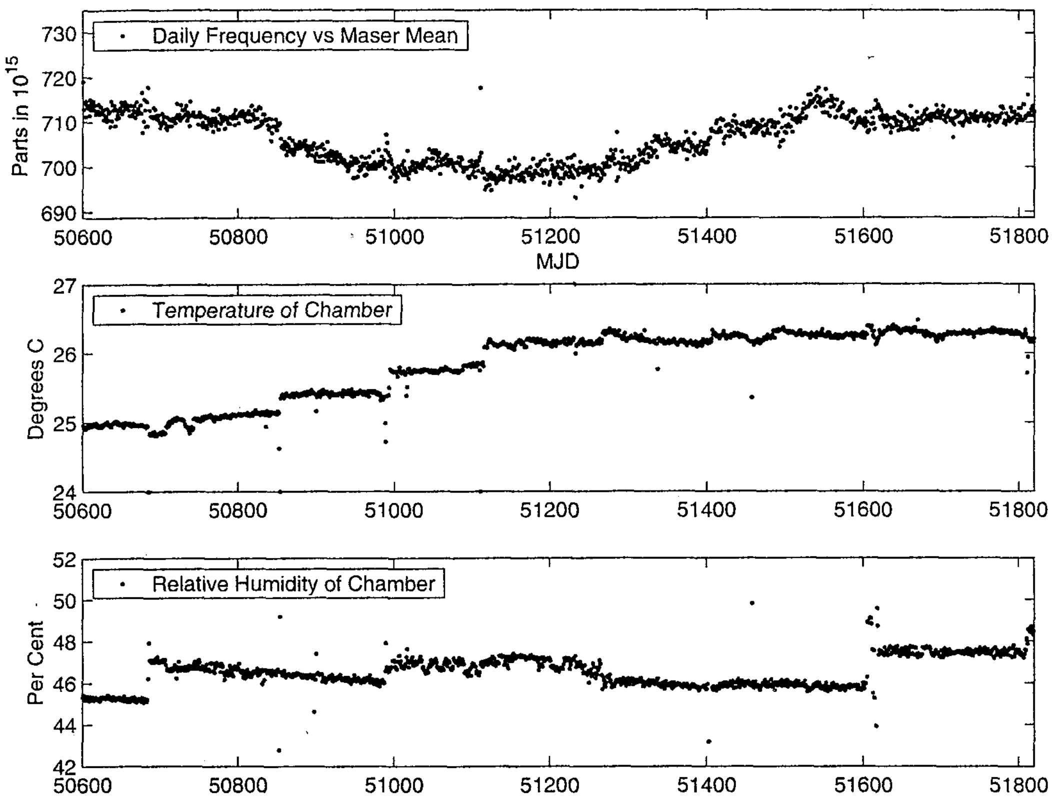


Figure 12. Observed frequency dependence of maser NAV3 on temperature and humidity as measured at the chamber ceiling.

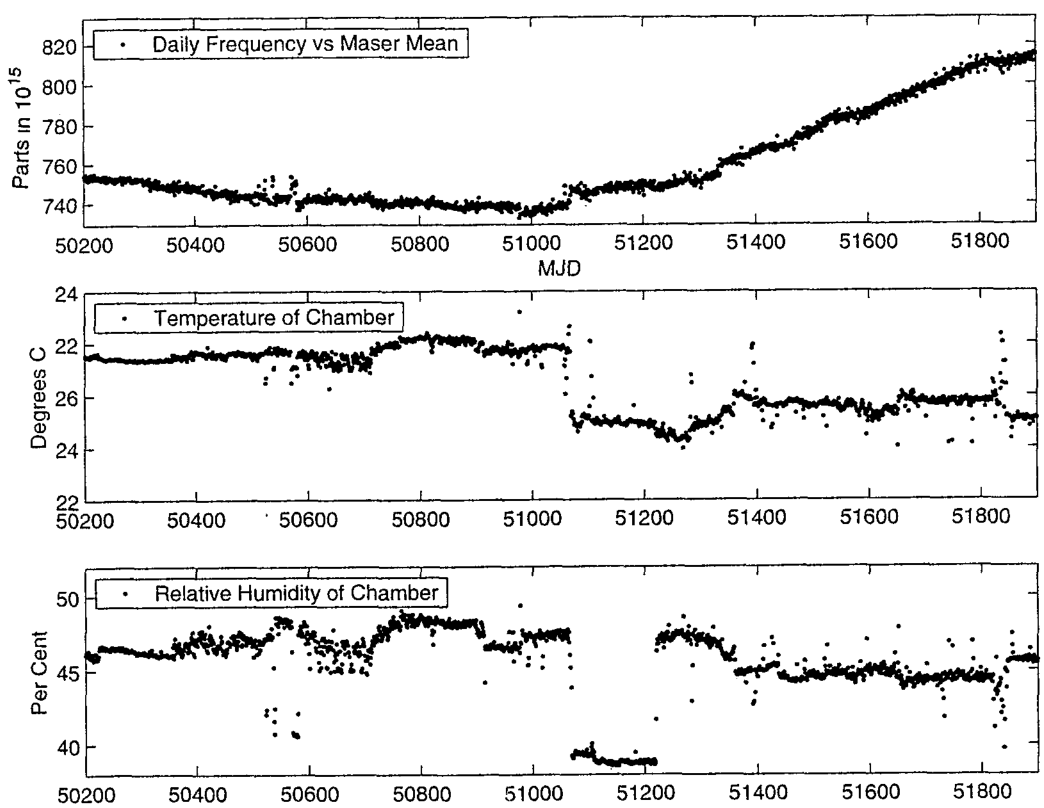


Figure 13. Observed frequency drift dependence of maser NAV8 on temperature and humidity as measured at the chamber ceiling.

Questions and Answers

DAVID HOWE (NIST): One of the comments you made, Lee, was that you take 60 days worth of data to get an estimate of the drift and correct for that drift. Yet, at the same time, you admit that the drift changes in those hydrogen masers for shorter intervals.

LEE BREAKIRON: Shorter than 60 days? No.

HOWE: You commented that you could make corrections more frequently, but that would raise the level of random walk. Do I understand that correctly?

BREAKIRON: Yes.

DAVE HOWE: Okay, then what I would comment about is that as your administered effort to reduce the random walk invariably will result in an increase in the noise on the drift in longer term. In other words, you will stretch the instabilities out farther. Now, we've observed this at NIST. One of the reasons why it is a dangerous business to have hydrogen masers without sufficiently periodic frequency evaluations is that you cannot predict an event in hydrogen maser reliably in terms of its drift. The drift is high. I believe that the Hadamard will report numbers that are significantly higher than one would expect.

HAROLD CHADSEY (USNO): I am interested, Lee, in two items. First of all, you said you did this analysis over 6 years, and you said using our current method of data analysis. As a person doing the real-time analysis, I am curious to know what standards you used, because we have changed them over the past 3e years. In fact, we have changed them over the past 6 months as to what we are evaluating as being good, bad, and mediocre clocks. I am also interested in knowing in your paper, are you going to have the specific criteria for Set #1, #2, #3, and #4 so we can figure out exactly what is the best and see if we can do that in real time versus the postprocessing that you did?

BREAKIRON: Certainly the latter we can do. I quoted the numbers for the cesium masers for Set #3, which was 5 parts in 10^{15} for a significant rate change for a cesium, and a change of 5 parts in 10^{18}/day for the drift of a maser.

As far as the change in the operational criteria over time, I found those to be much less significant than the rate and drift changes and the differences between the different sets of criteria. So I don't think that will be an important effect. But we will publish those numbers.

THOMAS CLARK (NASA Goddard Space Flight Center): I'm interested, since you have sort of the longest span and operation on the largest number of clocks of both types, that you now have talked more definitively in statistics. Certainly, from the masers that we run in VLBI, I have some feeling for the answer to this question. That is, in VLBI, in the masers, we find that vac ion pumps are the most common failure and that tends to be about 3 to 4 years that we see vac ion pump problems. Can you comment on what your experiences are on how long standards will keep running before the experience of a problem. And then the other question I was going to ask, which is unrelated, are you starting to include any fountain data in these ensembles and how is it performing?

BREAKIRON: Well, we've done only minimal adjustments to our masers, changes of heating plates and batteries, I think. But no large items, though that will presumably become a problem at some point. And perhaps one of our engineers here can tell you more.

As far as the fountain data, no that is still an experimental device. Bu, we certainly look forward to using the data when they become available.

CHADSEY: To answer your question, Tom, it depends on the manufacturer of the maser as to what pressures they are running and things like that as to when we are seeing failures. The Sigma Tau masers that we are running have been running very well, with minor things such as battery replacements. The SAO masers are running at higher pressures. We were doing plates and glassware on them. About 3 to 4 years was what we're getting out of them. But they are running significantly higher pressures in the chambers, so you would expect that.

DEMETRIOS MATSAKIS (USNO): Let me add another comment. A lot of the errors that we see with the masers are not necessarily intrinsic to the masers. We keep them in chambers that nominally are good to a 0.1°C. But every now and then, there is an adjustment to the chambers for one reason or another. The whole chamber can vary by half a degree, or on that level, and reach a new set point because of things we have done. We see that in the frequency of the masers. So part of the errors in the masers would not be there if they had been in a pristine, really perfect temperature-controlled environment. They are, however, in the best that we can do.